Education, Frisch Elasticity and Incentives

Preliminary and Incomplete*

Julio A. Blanco and Axelle Ferriere, New York University

December 14, 2013

Abstract

We study optimal insurance arrangement in a dynamic life-cycle Mirrlees economy. Labor productivity is the product of human capital and skills. Skills are private information and follow a stochastic Markov process. Human capital is an endogenous decision and could be either private or public information. We assume regularity conditions on the human capital technology in order to reduce the state space and make the problem tractable. We provide analytical characterization of the labor wedge and its life cycle behavior. Also, when human capital is public information, we provide a simple formula relating the Euler equation for human capital and the labor wedge. We compute the solution for a three period example.

*Preliminary and incomplete, comments are very welcomed. Email: aaf287@nyu.edu, jab772@nyu.edu.

1 Introduction

There are two main characteristics for wages over the life-cycle: first, they increase over time, second, they are subject to random variations. Figure 1 describes these two facts, plotting mean and variance of wages over the life-cycle. This evidence motivates this paper to analyze optimal insurance arrangements, when wages have two components: (1) an exogenous, stochastic, and persistent component, that we define as skills and (2) an endogenous component, human capital. Human capital accumulation generates the increasing pattern for mean wages while stochastic skills determines the variance.

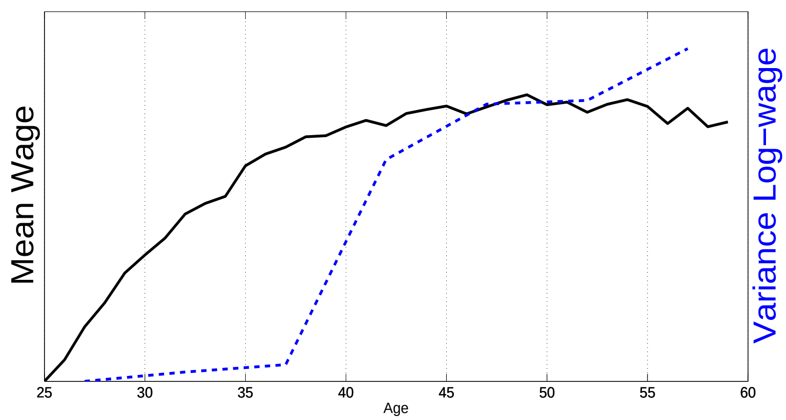

Figure 1: Dynamic of mean and variance of wages.

The main assumption throughout the paper is that the decomposition between wages and labor supply is unobservable and therefore optimal insurance agreements must respect incentives, in a Mirleesian fashion. Concerning human capital, we work with two sets of assumption. First, human capital is assumed observable but skills are not (adverse selection). Then, both human capital and skills are assumed private information to worker (adverse selection and moral hazard). The goal of this paper is to answer the following questions. What is the optimal trade-off between efficiency and insurance when workers accumulate human capital? Reciprocally, how do optimal insurance schemes affect incentives over education?

As a benchmark, we first study the full-information case, where both efforts and skills are observable every period. We assume that utility is separable between consumption and leisure, and that Frisch elasticity is constant over total time spent, in education and on the labor market. We characterize the first-best optimal level of effort spent in human capital. Because effort decreases over the life cycle while hours worked increase, the Frisch elasticity decreases over the life-cycle in average. As an implication, wages are U-shaped over the life-cycle, in line with the data. The decreasing pattern of the labor-market Frisch elasticity appears key for optimal insurance agreements, as it will shape incentives.

We turn to an economy where time spent in education is observable, but skills shocks and time spent on the labor market are not. To make the problem tractable, we impose a simple restriction on human capital accumulation technology, relating the persistence of skills to human capital depreciation. With this assumption, we only need to keep track of *total labor productivity*, that is, the product of human capital and skills, rather than tracking the composition of labor productivity. We show that the Inverse Euler equation holds. We obtain a dynamic characterization of the labor wedge as a function of two components: the life-cycle profile of the Frisch elasticity and the accumulated effect of incentives over time. Both elements generate an increasing dynamic for the labor wedge. The former force is given by the decreasing cost of incentives when the Frisch elasticity decreases (it is optimal to tax less elastic goods). The latter conveys the accumulation of persistent incentives over the life cycle. We also characterize the optimal human capital accumulation. As compared to the perfect information case, investment in education is affected by taxes, that reduce the return on human capital. On top of that, there is an additional force that captures how time spent in education affects incentives *in the present and in the future*. The dynamic cost of incentives generates an extra force in human capital accumulation.

Finally, we study the economy with the double source of private information: adverse selection (skills) and moral hazard (effort). We use a first-order approach for both incentives and numerically check for double deviations. We show that ...

Section 2 analyze the contribution of this paper to the literature. Section 3 present the environment. Section 4, 5 and 6 characterize the optimal allocation with public information, private information in skills, and finally private information in both skills and human capital. Section 7 presents numerical results and Section 8 concludes.

2 Related Literature

This paper contributes to the literature of optimal taxation based on dynamic private information in a life-cycle model with human capital. We divide the study of optimal taxation in four blocks: redistribution vs incentives, with and without human capital, and insurance vs incentives, again with and without human capital. By insurance, we mean that agents face uncertainty along their life-cycle: the Planner partly insures agents against this (individual) uncertainty. On the other hand, we think of redistribution as insurance

	Human Capital	
	Without	With
Redistribution	Mirrlees (1971) Saez (2001)	Kapicka (2006), Bohacek and Kapicka (2008) Kapicka (2011)
Insurance	**?**, Golosov, Troshkin, and Tsyvinski (2011), **?**	Grochulski and Piskorski (2010) **?**, Blanco and Ferriere (2013)

Table 1: Related Literature

provided by the Planner across agents against an initial and constant-across-life shock (productivity when born for instance). Table 1 describes the four blocks.

The basic proposal of Mirrlees (1971) consisted in studying how a government can design taxes when they do not observe wages (or other variables reflecting marginal productivity) but can observe labor income. The original Mirleesian model is static: marginal productivity varies across agents but is fixed per agent. In this framework, the author proves that this informational constraint generates a wedge between the marginal benefit and marginal cost of leisure, and shows how to construct an optimal tax system that minimizes this distortion. Together with more recent related papers[1], we have a clear idea of the impact of taxes on welfare, and of how to construct optimal taxes when the government is constrained in his information about marginal productivities or salaries among individuals.

Nevertheless, empirical evidence suggests that marginal productivity of labor fluctuates over time, but with persistence, at individual levels. This motivates the original paper of Golosov, Kocherlakota, and Tsyvinski (2003), which studies how unobservable persistent productivity shocks affect capital on taxes on the top of labor income. More recent papers by Farhi and Werning (2011) and Golosov, Troshkin, and Tsyvinski (2011) take the quantitative step of applying optimal taxation in a life-cycle model, where the trend on labor productivity is entirely exogenous. In these models, the efficiency vs. insurance trade-off typically generates a path for labor taxes where both mean and variance of taxes increase with age.

On the other hand, Kapicka (2006), Bohacek and Kapicka (2008) and Kapicka (2011) extend the original static Mirleesian model with fixed productivity to a dynamic model with fixed productivity and human capital accumulation. Bohacek and Kapicka (2008) develop a model where agents have an unobservable idiosyncratic productivity component that is fixed over the life-cycle, and accumulate observable human capital every period. Kapicka (2011) extend the model to unobservable human capital accumulation, and show that the average labor wedge typically decreases over the life-cycle. This is because labor wedges at the end of the life-cycle decrease returns of human capital investment in every period of the life-cycle, and

[1]See for instance Saez (2001).

are therefore much more distorting than wedges earlier in life.

Some papers have started extending these Mirleesian models with human capital, where agents face fluctuating productivity shocks over time. Grochulski and Piskorski (2010) study a problem with unobservable (risky) human capital, but schooling investment happens only in initial period: there is no interaction between life-cycle uncertainty and investment in human capital over the life-cycle in their framework. Finally, **?** extend the model of Kapicka (2011) to varying productivity shocks. Their set-up is therefore close to our model. In fact, their theoretical results are in line with our findings: the "no distortion at the top" result may not hold in this richer framework, and the average labor wedge may increase or decrease over the life-cycle (XXX). However, **?** numerically solve their model only for two periods: without the assumption on the depreciation rate of human capital, the model is not numerically tractable.

Therefore, our contribution relies in the study of both human capital and persistent productivity shocks in a tractable private information economy. In particular, the simple assumption on the depreciation rate of human capital allows us to get a tractable framework, and therefore to quantitatively compute the average mean and variance for labor taxes over the life-cycle, when adverse selection and moral hazard problems are combined.

3 Model

3.1 Preferences and Technology

Time is discrete and finite: $t = 0, 1, 2, \ldots, T$. Assume a continuum of agents of measure one, indexed by s. Agents have preferences given by:

$$U(c, l, e) = \mathbb{E}\left[\sum_{t=0}^{\infty} u(c_t, l_t, e_t)\right] \tag{1}$$

where c, l, e stand for sequences of consumption, labor supply in output and labor supply in human capital accumulation. We call l_t labor and e_t education in the rest of the paper. We assume standard properties for the period utility function: $u(\cdot)$ is increasing in c, decreasing in l and e, and concave in all arguments. Output in period t is generated by a linear technology over labor given by:

$$y_t = h_t \theta_t l_t \tag{2}$$

where y_t is output, and (h_t, θ_t, l_t) stand for idiosyncratic human capital, exogenous productivity and labor in period t. The technology for human capital accumulation is given by:

$$h_t = G_t(h_{t-1}, e_t) \tag{3}$$

where $G_t(\cdot, \cdot)$ is increasing and concave in both arguments. Note that *the effort chooses in period t generates human capital that is productive in period t.* Individual productivity follows a Markov process described by the transition probability $f_t(\theta_{t+1}|\theta_t)$. Let F_t denote the cumulative density of f_t.

Finally, it is standard to redefine preferences as a function of net output y instead of labor:

$$U(c, y, e, h, \theta) = \mathbb{E}\left[\sum_{t=0}^{\infty} u(c_t, \frac{y_t}{\theta_t h_t}, e_t)\right] \tag{4}$$

3.2 Information and Environment

Let θ^t be the sequence of productivity shocks: $\theta^t \equiv (\theta_{-1}, \theta_0, \theta_1, \ldots, \theta_t)$ and $\hat{\theta}^t$ the history of reports. We denote each individual with a history of shocks θ^t. For simplicity we assume that θ_{-1} and h_{-1} are equalized across agents. We define contracts such that output and consumption in period t are function of the history of reports up to time t, and e_t is a function of the history of reports up to time $t-1$, therefore

$$\{c_t(\hat{\theta}^t), y_t(\hat{\theta}^t), e_t(\hat{\theta}^{t-1}), h_t(\hat{\theta}^{t-1})\} \tag{5}$$

It is important to notice the assumption that e_t and h_t are θ^{t-1}-measurable, i.e. are not affected by the time-t realization of uncertainty. This assumption captures the risky nature of human capital investment: without knowing the return on human capital at time t, the planner has to choose human capital in that period.

We characterize efficient allocation under three scenarios:

1. **Perfect Information**: This is the benchmark case in all variables are public information $(c_t, y_t, l_t, h_t, e_t, \theta_t)$, in particular θ_t and e_t (and then h_t) are observable every period.

2. **Imperfect Information over exogenous labor productivity**: In this case (c_t, y_t, h_t, e_t) are public information and θ_t together with l_t is private information to the agent.

3. **Imperfect Information over total labor productivity**: In this case (c_t, y_t) are public information and all other variables are private information.

The planning problem is given by

$$K(v) = \min_{c,y,h,e} \mathbb{E}\left[\sum_{t=0}^{\infty} q^t \left(c_t(\theta^t) - y_t(\theta^t)\right) |\theta_0\right] \tag{6}$$

s.t.

$$\begin{aligned} v &= U(c,y,e,h,\theta) && (7)\\ h_{t+1}(\theta^t) &= G_{t+1}(h_t(\theta^{t-1}), e_{t+1}(\theta^t)) \;\; \forall t \; \forall \theta^t && (8)\\ \{c,y,h,e\} & \quad \text{is incentive compatible} && (9)\end{aligned}$$

The last constraint is given by the information structure. In the case of perfect information, the IC constraint is empty. We will be more explicit about the nature of the IC in the two other scenarios. Finally, throughout the paper we follow two key assumptions to characterize the problem. First, we assume additive preferences between consumption and time - allocated both on the labor market and in education[2]. The second assumption imposes some restrictions on the stochastic behavior of exogenous productivity together with human capital production technology.

Assumption 1 *$u(c, \frac{y}{\theta h}, e) = u(c) - h(\frac{y}{\theta h} + \xi e)$, with $\xi > 0$.*

Assumption 2 *Let $\varphi_t^1(x), \varphi_t^2(y)$ a sequence of functions s.t.*

$$\varphi_t^1(x).\varphi_t^2(y) = \varphi_t(xy) \tag{10}$$

Then $f_t(\theta_{t+1}|\theta_t)$ satisfies

$$Pr(\theta_{t+1} \leq c|\theta_t) = F_t\left(\frac{c}{\varphi_t^1(\theta_t)}\right) \tag{11}$$

for all θ_t, and there exists a function $H_{t+1}(\cdot)$ such that $G_t(\cdot)$ satisfies

$$G_{t+1}(h_t, e_{t+1}) = \varphi_t^2(h_t) H_{t+1}(e_{t+1}) \tag{12}$$

all t, h_t, e_{t+1} and θ_t.

The latter assumption holds when the depreciation rate of $\log h$ is equal to one minus the persistence

[2] For some of our results, this assumption is not necessary. However, to keep the paper simple, we maintain this assumption throughout the paper.

rate of $\log\theta$: both exogenous productivity θ and endogenous productivity h depreciate at the same rate. To clarify this assumption, we describe two examples where this assumption does and does not hold.

Example 1 *Assume that* $\theta_t \sim_{i.i.d.} \epsilon_t$ *and* $h_{t+1} = (1-\delta)h_t + (h_t e_{t+1})^\alpha$. *Or, assume that* $\theta_{t+1} = \theta_t^\rho \epsilon_{t+1}$ *and* $h_{t+1} = e_{t+1}$. *Then, Assumption 2 does not hold.*

Example 2 *Assume that* $\theta_{t+1} = \theta_t^\rho \epsilon_{t+1}$ *and* $h_{t+1} = Ah_t^\rho(1 + (e_{t+1})^\alpha)$. *Then, Assumption 2 holds.*

The key difference between the two examples is the "persistence" in both variables. In the first case one variable is contemporaneous while the other depends on the past, implying that total labor productivity θh is not Markovian - the composition of labor productivity matters to predict future labor productivity. In the second example, total labor productivity is Markovian. What is the empirical validity of this assumption?

Both idiosyncratic productivity shocks and human capital are difficult to measure empirically. Depending on the model, estimations of the persistence rate of shocks on individual skills vary significantly. Using PSID data on real wage across households, Storesletten, Telmer, and Yaron (2004) find a persistence rate larger than 0.95. When inferring individual skills from a model with private information, Golosov, Kocherlakota, and Tsyvinski (2003) find numbers closer to 0.8. When explicitly modeling human capital, Yaron (2011) estimate a persistence rate of shocks - to human capital in their model - also around 0.8. This therefore pins down a depreciation rate for human capital that is between 5% and 20% per period. In our model, θ and h enter both the production function and the total productivity of each agent in the same way, and we argue that assuming an equal rate of depreciation for these two variables is a reasonable approximation.

4 Perfect Information

In this section, we characterize the allocation with perfect information. The objective is threefold. First, a neat characterization of the recursive problem is proposed. Second, consumption, labor and human capital allocations are described. Third, we characterize the trend in labor earning generated by human capital accumulation, and show how this trend translates in a time-varying Frisch elasticity of labor supply.

The first theorem states that the optimal allocation can be (recursively) characterized as a function of the total labor productivity $\tilde{\theta} \equiv \theta h$, Markovian by Assumption 2.

Proposition 4.1 *Let Assumption 2 hold. Let* $K(\tilde{\theta}, v, t)$ *the present discount value of excess of expenditures* *at time t with a promise value v and labor productivity $\tilde{\theta}_t$, then $K(\tilde{\theta}, v, t)$ satisfies*

$$K(v, \tilde{\theta}_{-1}, t) = \min_{\tilde{c}(\tilde{\theta}), \tilde{y}(\tilde{\theta}), \tilde{v}(\tilde{\theta}), e} \{ \int_{\Theta} \tilde{c}(\tilde{\theta}) - \tilde{y}(\tilde{\theta}) + q\tilde{K}(v(\tilde{\theta}), \tilde{\theta}, t+1) \tilde{f}_t(\tilde{\theta}|\tilde{\theta}_{-1}, e) d\tilde{\theta} \} \tag{13}$$

with

$$\tilde{w}(\tilde{\theta}) = U(\tilde{c}(\tilde{\theta}), \frac{\tilde{y}(\tilde{\theta})}{\tilde{\theta}}, e) + \beta \tilde{v}(\tilde{\theta}) \tag{14}$$

$$v = \int_{\Theta} \tilde{w}(\tilde{\theta}) \tilde{f}_t(\tilde{\theta}|\tilde{\theta}_{-1}, e) d\tilde{\theta} \tag{15}$$

$$\tilde{f}_t(\tilde{\theta}|\tilde{\theta}_{-1}, e) = \frac{f_t\left(\frac{\tilde{\theta}}{\varphi_t(\tilde{\theta}_t) H_t(e)}\right)}{\varphi_t(\tilde{\theta}_t) H_t(e)} \tag{16}$$

Next, we characterize the allocations under perfect information, with a key emphasis on the life-cycle profile of net output and education, since this will determine Frisch elasticity. Given that there is no Inada Condition on the labor supply, endogenous entry in the labor market will be determined by the calibration.

Proposition 4.2 *Let Assumptions 1 and 2 hold. Assume that $h(x) = \frac{1}{\alpha} x^{\alpha}$ and $q = \beta$, then the optimal effort for human capital allocation satisfies:*

$$\chi_t(e_t, \tilde{\theta}_{t-1}) = \frac{\frac{dH_t(e_t)}{de}}{H_t(e_t)} \left[\mathbb{E}[y(\tilde{\theta}_t)|\tilde{\theta}_{t-1}, e_t] + q\mathbb{E}[\tilde{\theta}_t \chi_{t+1}(e_{t+1}, \tilde{\theta}_t) \frac{\frac{d\varphi_{t+1}(\tilde{\theta}_t)}{d\tilde{\theta}'}}{\varphi_{t+1}(\tilde{\theta})} \frac{H(e_{t+1})}{\frac{dH(e_{t+1})}{de}} |\tilde{\theta}_{t-1}, e_t] \right] \tag{17}$$

where,

$$\chi(e_t, \tilde{\theta}_{t-1}) \equiv \xi \left(\mathbb{E}[\tilde{\theta}_t > \tilde{\theta}^{\star}_{t-1}|\tilde{\theta}_{t-1}, e_t](1 - F_t(\tilde{\theta}^{\star}_{t-1}|\tilde{\theta}_{t-1}, e_t)) + \theta^{\star}_{t-1} F_t(\tilde{\theta}^{\star}_{t-1}|\tilde{\theta}_{t-1}, e_t) \right) \tag{18}$$

$$\tilde{\theta}^{\star}_{t-1} \equiv (\xi e_t)^{\alpha-1} u_c(c(v_0))^{-1} \tag{19}$$

$$y(\tilde{\theta}_t) = \tilde{\theta}_t \left[\left(u_c(c(v_0)) \tilde{\theta}_t \right)^{\frac{1}{\alpha-1}} - \xi e_t \right] \; \forall \tilde{\theta}_t \geq \tilde{\theta}^{\star}_{t-1}, \quad 0 \textit{ otherwise} \tag{20}$$

$$\tilde{\theta}_t = \varphi_t(\tilde{\theta}_{t-1}) H_t(e_t) \epsilon_t \tag{21}$$

Note that $\tilde{\theta}^{\star}_{t-1}(\tilde{\theta}_{t-1}, e_t)$ defines the threshold such that a $\tilde{\theta}_{t-1}$-worker that receives productivity $\tilde{\theta}_t \geq \tilde{\theta}^{\star}_{t-1}$ works in period t. Below this value, $y(\tilde{\theta}_t) = 0$.

The characterization for consumption and hours of work are usual in the literature: with perfect information, the optimal allocation for consumption exhibits full-insurance. Consumption is history-independent, and constant when $q = \beta$. The assumption that time spent on the labor market and in human capital are perfect substitutes gives a sharp characterization of effort as a function of labor productivity as Equation 17 shows: marginal cost of capital accumulation, on the left-hand side, is equal to marginal benefit, on the right-hand side. The marginal cost, $\chi_t(e, \theta)$, captures the cost of decreasing labor supply, if labor supply is positive. The marginal benefit is given by the benefit of increasing productivity from period t onward: the first term reflects the increase in output in period t, and the second term, the marginal benefit of reducing the effort tomorrow, keeping labor productivity constant.

Finally, we simulate the economy for 34 periods, using the following laws of motion for productivity and human capital accumulation given by:

$$\theta_{t+1} = \theta_t^{\rho} e^{\sigma_\epsilon \epsilon_{t+1} - \sigma_\epsilon^2}$$

$$h_{t+1} = a_0 h_t^{\rho}(1 + a_1 e_t^{\gamma})$$

where $\epsilon_{t+1} \sim \mathcal{N}(0, 1)$. We roughly calibrate the model to life-cycle mean and variance wages in the US.[3]

The first important observation is that the restriction on the law of motion for human capital imposed by Assumption 2 does not prevent us to match the life-cycle mean profile of wages. The top panels of Figure 2 plots the life-cycle mean and standard deviation of wages in our model and in the data: the fit is reasonable. The second implication relates to the Frisch elasticity of labor supply: in our model, while the Frisch elasticity of total time spent in labor *plus* education is constant (equal to unity) due to our utility function, the Frisch elasticity of labor *only* moves over the life cycle, as plotted in the bottom left panel of Figure 2. By definition, the Frisch elasticity of labor supply η_t is given by:

$$\eta_t \equiv \frac{1}{\alpha - 1} \frac{l_t + \xi e_t}{l_t} \tag{22}$$

The life-cycle profile in the allocation of time - between labor and education - implies a life-cycle pattern for the Frisch elasticity as well. The bottom right panel of Figure 2 shows the life-cycle mean profiles of education and labor. Due to persistence in human capital, it is optimal to accumulate human capital at the beginning of the life-cycle: education e_t is decreasing over the years, converging to zero at the end of the working life. On the other hand, because productivity $\tilde{\theta}$ increases over the life-cycle, labor supply also increases. Thus, the Frisch elasticity exhibits a decreasing life-cycle profile, converging to unity at the end of the life-cycle.

[3]We simulate ten thousand workers. We use $u(c) = \log c$, $\beta = q = 0.96$, $\alpha = 2$ and $\xi = 1$ in the utility function, $\rho = 0.99$, $\sigma_\epsilon = 0.0326$, $a_0 = 0.99$, $a_1 = 0.055$ and $\gamma = 0.5$ for productivity and human capital laws of motion.

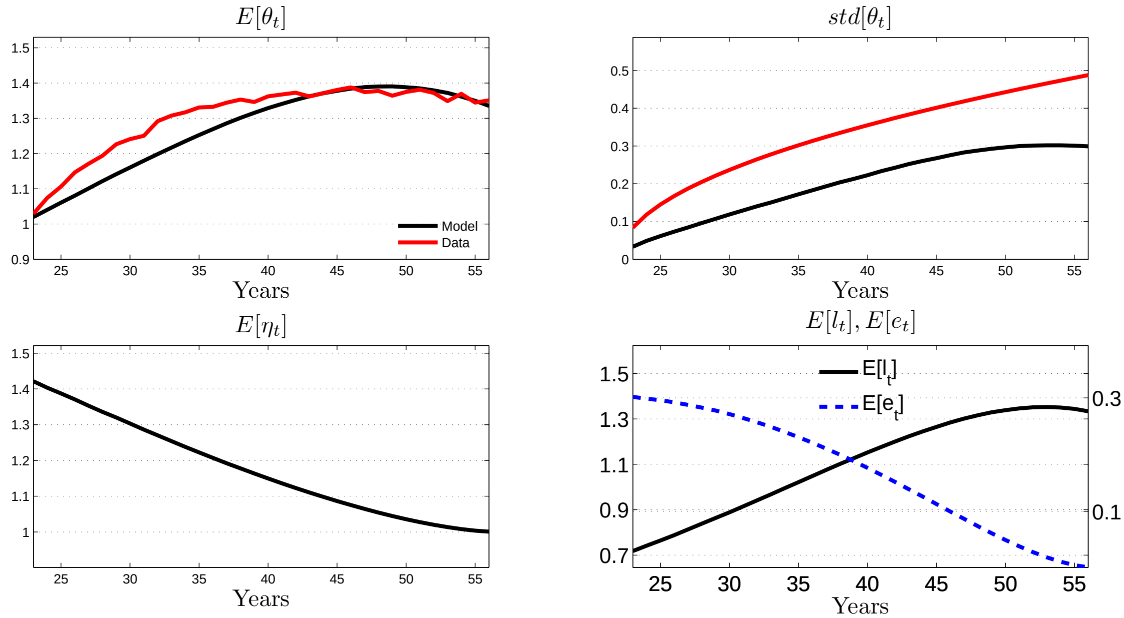

Figure 2: Life Cycle Profile of Wages: Data and Model

5 Imperfect Information over exogenous labor productivity

We turn to the problem with adverse selection, when skills and time spent on the labor market are not observable, but education and human capital are. The next proposition characterizes the recursive problem.

Proposition 5.1 *Assume [2]. Let $K(\tilde{\theta}, v, t)$ the discounted present value of excess of expenditures at time t with promised value v and labor productivity $\tilde{\theta}_t$, then $K(\tilde{\theta}, v, t)$ satisfies*

$$K(v, \Delta, \tilde{\theta}_{-1}, t) = \max_{c(\tilde{\theta}), y(\tilde{\theta}), v(\tilde{\theta}), \Delta(\tilde{\theta}), e} \{ \int_{\Theta} c(\tilde{\theta}) - y(\tilde{\theta}) + qK(v(\tilde{\theta}), \Delta(\tilde{\theta}), \tilde{\theta}, t+1) f_t(\tilde{\theta}|\tilde{\theta}_{-1}, e) d\tilde{\theta} \} \tag{23}$$

with

$$w(\tilde{\theta}) = U(c(\tilde{\theta}), \frac{y(\tilde{\theta})}{\tilde{\theta}}, e) + \beta v(\tilde{\theta}) \tag{24}$$

$$\dot{w}(\tilde{\theta}) = U'_{\tilde{\theta}}(c(\tilde{\theta}), \frac{y(\tilde{\theta})}{\tilde{\theta}}, e) + \beta \Delta(\tilde{\theta}) \tag{25}$$

$$v = \int_\Theta \tilde{w}(\tilde{\theta}) \tilde{f}_t(\tilde{\theta}|\tilde{\theta}_{-1}, e) d\tilde{\theta} \quad (26)$$

$$\Delta = \int_\Theta \tilde{w}(\tilde{\theta}) \frac{\partial \tilde{f}_t(\tilde{\theta}|\tilde{\theta}_{-1}, e)}{\partial \tilde{\theta}} d\tilde{\theta} \quad (27)$$

and

$$f_t(\tilde{\theta}|\tilde{\theta}_-, e) = \frac{f\left(\frac{\tilde{\theta}_t}{\varphi(\theta_-)H_t(e)}\right)}{\varphi(\theta_-)H_t(\tilde{\theta}_t, e)} \quad (28)$$

Why do incentives not break the result for the IC constraint, when human capital is observable? Does this result depend on the use of the first-order approach? The formal answer to these two questions is in the appendix but it is insightful to highlight the intuition for the last period, T. A necessary condition for the global IC to hold is that the local IC holds. Let $w(\theta^T)$ the utility of the contract under truth-telling after history θ^T. If the contract satisfies the IC constraint then

$$w(\theta^T) = \max_r \{u(c(\theta^{t-1}, r), \frac{y(\theta^{t-1}, r)}{\theta_t h_t(\theta^{t-1})}, e(\theta^{t-1}))\}$$

Using the first-order approach and dividing both sizes by $\frac{1}{h(\theta^{t-1})}$

$$\begin{aligned} \frac{\partial w(\theta^t)}{\partial \theta_t} \frac{1}{h(\theta^{t-1})} &= \frac{1}{h(\theta^{t-1})} \left[u'_\theta(c(\theta^{t-1}, r), \frac{y(\theta^{t-1}, r)}{\theta_t h_t(\theta^{t-1})}, e(\theta^{t-1})) \right] \\ &= u'_l(c(\theta^{t-1}, r), \frac{y(\theta^{t-1}, r)}{\theta_t h_t(\theta^{t-1})}, e(\theta^{t-1})) \frac{y(\theta)}{(h(\theta^{t-1})\theta)^2} \end{aligned}$$

Note that the right-hand size only depend on *total* labor productivity only, and not on the composition of labor productivity. Therefore knowing *total labor productivity* **only** is enough to generate a contract where the Planner can separate different agents. As before, given that total labor productivity is markovian, we can repeat this static argument to a dynamic context.

Next we characterize optimal allocations and wedges in this problem. The first Proposition states the standard inverse Euler Equation. Note, however, that with human capital accumulation, time spent in education e_{t+1} determines the expectation $\mathbb{E}_t$, as it is θ_t-measurable.

Proposition 5.2 *The consumption satisfies the inverse Euler equation:*

$$u'(c_t)^{-1} = \frac{q}{\beta} \mathbb{E}_t[u'(c_{t+1})^{-1}] \quad (29)$$

To characterize the labor wedge, and through the rest of the paper, we use the following function for h:

$$h(y_t, e_t; \tilde{\theta}_t) = \frac{\kappa}{\alpha}\left(\frac{y_t}{\tilde{\theta}_t} + \xi e_t\right)^{\alpha} \tag{30}$$

Proposition 5.3 *Let τ^L be the labor wedge defined as $\tau_t^L(\theta^t) \equiv 1 - h_y(y_t(\theta^t), e_t(\theta^t); \theta^t)/u_c(c(\theta^t))$. Let $\pi(\tilde{\theta})$ be a integrable function and define $\Pi(\theta)$ the primitive of $\frac{\pi(\tilde{\theta})}{\tilde{\theta}}$. If the optimal solution of labor supply is interior, then the labor wedge satisfies*

$$\mathbb{E}_{t-1}\left[\frac{\eta_t}{1+\eta_t}\frac{\tau_t^L}{1-\tau_t^L}\frac{q}{\beta u'(c_t)}\pi_t\right] = \mathbb{E}_{t-1}\left[\Pi_t\left(\frac{q}{\beta u'(c_t)} - \frac{1}{u'(c_{t-1})}\right)\right] + \frac{\eta_{t-1}}{1+\eta_{t-1}}\frac{\tau_{t-1}^L}{1-\tau_{t-1}^L}\frac{\tilde{\theta}_{t-1}}{u'(c_{t-1})}\mathbb{E}_{t-1}\left[\Pi_t L_t\right] \tag{31}$$

where

$$L_t = \frac{\frac{\partial f_t(\tilde{\theta}_t|\tilde{\theta}_{t-1}, e_t)}{\partial \tilde{\theta}_{t-1}}}{f_t(\tilde{\theta}_t|\tilde{\theta}_{t-1}, e_t)} \tag{32}$$

Note that with $e_t \to 0$, $\eta_t = (\alpha-1)^{-1}\forall t$ and Equation 31 reverts to the labor wedge characterization described by Farhi and Werning (2011). To see more clearly the main property of the labor wedge, let us use $\varphi_t(x) = x^{\rho}$. Equation 31 becomes

$$\mathbb{E}_{t-1}\left[\frac{\eta_t}{1+\eta_t}\frac{\tau_t^L}{1-\tau_t^L}\frac{qu'(c_{t-1})}{\beta u'(c_t)}\right] = \rho\frac{\eta_{t-1}}{1+\eta_{t-1}}\frac{\tau_{t-1}^L}{1-\tau_{t-1}^L} + \mathbb{C}ov_{t-1}\left[\log\theta_t, \frac{qu'(u'(c_{t-1}))}{\beta u'(c_t)}\right] \tag{33}$$

How does this equation change with the dynamic behavior of the Frisch elasticity? First, the optimal labor wedge is affected by the persistence of the stochastic process for exogenous labor productivity: since labor productivity is persistent, distortions in the marginal rate of substitution due to incentives are also spread over time. The following term, that is, a combination of taxes and Frisch elasticity, therefore exhibits persistence ρ

$$\frac{\eta_t}{1+\eta_t}\frac{\tau_t^L}{1-\tau_t^L} \tag{34}$$

Then, two forces drive the labor wedge up in our model. First, as in Farhi and Werning (2011), the term $\frac{\eta}{1+\eta}\frac{\tau_L}{1-\tau_L}$ is increasing over age, because the covariance of consumption growth with the log of productivity is itself positive, in order to provide incentives. Second, given an increase in $\frac{\eta}{1+\eta}\frac{\tau_L}{1-\tau_L}$, the Frisch elasticity itself is decreasing over the life cycle. This generates an increasing trend for labor taxes across the life-cycle (efficiency) together with an extra volatility given by the contingent distribution and allocation of time in labor supply and human capital accumulation. Indeed, because, early in the life-cycle, a worker main goal is to accumulate human capital, a marginal increase in labor generate a large disutility lost and therefore taxation should be low. This is reminiscent of the standard result of taxing less more elastic goods. Note that we also obtain the typical result of zero taxation at the top and the bottom of the distribution of skills.

Proposition 5.4 $\tau_t^L(0) = 0$ *and* $\lim_{\tilde{\theta}\to\infty} \tau_t^L(\tilde{\theta}) = 0$.

Finally, we turn to the optimality conditions for human capital accumulation.

Proposition 5.5 *Assume 1 and 2. Then if labor supply is positive for all* $\tilde{\theta}_t, t \tilde{+} 1$ *the optimality condition for human capital is given by*

$$\chi_t(\theta_{t-1}, e_t) = \frac{H_t'(e_t)}{H_t(e_t)} \left[\mathbb{E} \left[\left(1 + \frac{\eta_t(\tilde{\theta}_t)\tau_t^L(\tilde{\theta}_t)}{1+\eta_t(\tilde{\theta}_t)} \right) y_t(\tilde{\theta}_t) | \theta_{t-1}, e_t \right] + q\mathbb{E} \left[\theta_t \chi_{t+1}(\theta_t, e_{t+1}) \frac{\varphi_{t+1}'(\tilde{\theta}_t) H_{t+1}(e_{t+1})}{\varphi_{t+1}(\tilde{\theta}_t) H_{t+1}'(e_{t+1})} | \theta_{t-1}, e_t \right] \right] + \dots$$

$$\dots + \frac{H_t'(e_t)}{H_t(e_t)} \mathbb{E} \left[\frac{\tau_t^L(\tilde{\theta}_t)}{1-\tau_t^L(\tilde{\theta}_t)} \frac{\tilde{\theta}_t \eta_t((\tilde{\theta}_t))}{1+\eta_t((\tilde{\theta}_t))} \left(\frac{\dot{w}_t(\tilde{\theta}_t)}{u_t'(c_t(\tilde{\theta}_t))} + q \frac{\theta_t d \left(\frac{\varphi_{t+1}'(\tilde{\theta}_t)}{\varphi_{t+1}(\tilde{\theta}_t)} \right)}{d\tilde{\theta}_t} \frac{u'(c_{t+1}(\tilde{\theta}_{t+1}))}{u'(c_t(\tilde{\theta}_T))} \theta_{t+1} \frac{\dot{w}_{t+1}(\tilde{\theta}_{t+1})}{u'(c_{t+1}(\tilde{\theta}_{t+1}))} \right) | \theta_{t-1}, e_t \right]$$

Where $\chi_t(\theta_{t-1}, e_t)$ *is equal to*

$$\chi_t(\theta_{t-1}, e_t) = \xi\mathbb{E} \left[\tilde{\theta}_t \left(1 - \frac{\eta_t(\tilde{\theta}_t)\tau_t^L(\tilde{\theta}_t)}{1+\eta_t(\tilde{\theta}_t)} \right) | \theta_{t-1}, e_t \right] \tag{35}$$

First of all, note that there is no distortion in human capital accumulation when $\tau_L = 0$: Equation 35 collapses to its perfect-information counterpart.[4] The left-hand side of Equation 35 captures the marginal cost of effort: it reflects the marginal benefit of labor supply, taking into account taxation. The first term of the right-hand side represents the marginal benefit, on output and effort next period, keeping constant the total output. Again, taxation enters this term. Finally, the last term of the right-hand side captures how the perturbation in effort directly affects incentives, today and tomorrow.

Corollary 3 *Assume now that* $\varphi(\theta) = \theta^\rho$. *We can simplify the last proposition:*

$$\chi_t(\theta_{t-1}, e_t) = \frac{H_t'(e_t)}{H_t(e_t)} \left[\mathbb{E} \left[\left(1 + \frac{\eta_t(\tilde{\theta}_t)\tau_t^L(\tilde{\theta}_t)}{1+\eta_t(\tilde{\theta}_t)} \right) y_t(\tilde{\theta}_t) | \theta_{t-1}, e_t \right] + q\rho\mathbb{E} \left[\chi_{t+1}(\theta_t, e_{t+1}) \frac{H_{t+1}(e_{t+1})}{H_{t+1}'(e_{t+1})} | \theta_{t-1}, e_t \right] + \dots \right.$$

$$\left. \dots + \mathbb{E} \left[\frac{\tau_t^L(\tilde{\theta}_t)}{1-\tau_t^L(\tilde{\theta}_t)} \frac{\tilde{\theta}_t \eta_t((\tilde{\theta}_t))}{1+\eta_t((\tilde{\theta}_t))} \left(\frac{\dot{w}_t(\tilde{\theta}_t)}{u_t'(c_t(\tilde{\theta}_t))} - \rho q \frac{u'(c_{t+1}(\tilde{\theta}_{t+1}))\tilde{\theta}_{t+1}}{u'(c_t(\tilde{\theta}_t))\tilde{\theta}_t} \frac{\dot{w}_{t+1}(\tilde{\theta}_{t+1})}{u'(c_{t+1}(\tilde{\theta}_{t+1}))} \right) | \theta_{t-1}, e_t \right] \right]$$

where χ_t *is defined as in Equation 35.*

[4] Assuming an interior solution for output.

6 Unobservable effort

Due to our formalization, we can extend the problem to unobservable effort, without extending the state space. We just need to add one extra constraint, to guarantee that e is optimal for the worker.

Proposition 6.1 *Assume [2]. Let $K(\tilde{\theta}, v, t)$ the discounted present value of excess of expenditures at time t with promised value v and labor productivity $\tilde{\theta}_t$, then $K(\tilde{\theta}, v, t)$ satisfies*

$$K(v, \Delta, \tilde{\theta}_{-1}, t) = \max_{c(\tilde{\theta}), y(\tilde{\theta}), v(\tilde{\theta}), \Delta(\tilde{\theta}), e} \{ \int_\Theta c(\tilde{\theta}) - y(\tilde{\theta}) + qK(v(\tilde{\theta}), \Delta(\tilde{\theta}), \tilde{\theta}, t+1) f_t(\tilde{\theta}|\tilde{\theta}_{-1}, e) d\tilde{\theta} \} \tag{36}$$

with

$$w(\tilde{\theta}) = U(c(\tilde{\theta}), \frac{y(\tilde{\theta})}{\tilde{\theta}}, e) + \beta v(\tilde{\theta}) \tag{37}$$

$$\dot{w}(\tilde{\theta}) = U'_{\tilde{\theta}}(c(\tilde{\theta}), \frac{y(\tilde{\theta})}{\tilde{\theta}}, e) + \beta \Delta(\tilde{\theta}) \tag{38}$$

$$v = \int_\Theta \tilde{w}(\tilde{\theta}) \tilde{f}_t(\tilde{\theta}|\tilde{\theta}_{-1}, e) d\tilde{\theta} \tag{39}$$

$$\Delta = \int_\Theta \tilde{w}(\tilde{\theta}) \frac{\partial \tilde{f}_t(\tilde{\theta}|\tilde{\theta}_{-1}, e)}{\partial \tilde{\theta}} d\tilde{\theta} \tag{40}$$

$$\Delta = -\frac{\varphi'_t(\theta_{t-1})}{\varphi_t(\theta_{t-1})} \frac{H_t(e_t)}{H'_t(e_t)} \int_\Theta U_e(c(\tilde{\theta}), \frac{y(\tilde{\theta})}{\tilde{\theta}}, e) \tilde{f}_t(\tilde{\theta}|\tilde{\theta}_{-1}, e) d\tilde{\theta} \tag{41}$$

and

$$f_t(\tilde{\theta}|\tilde{\theta}_-, e) = \frac{f\left(\frac{\tilde{\theta}_t}{\varphi(\theta_-)H_t(e)}\right)}{\varphi(\theta_-)H_t(\tilde{\theta}_t, e)} \tag{42}$$

Equation 41 implies that the effort chosen by the household is the one that maximizes its utility. Numerically, we will have to check for double-deviations.

7 Application for the US Economy

In this section we solve the model numerically for some parameters. We choose a time period of 1 year and therefore $\beta = q = 0.96$ and we simulate the model for three periods.[5] We set preference parameters $u(c) = \log(c)$ and $h(l+e) = \frac{\kappa}{\alpha}(l + \xi e)^\alpha$ with a Frisch elasticity equal to 1 and $\kappa = \xi = 1$. We follow Violante ... and calibrate $\theta_{t+1} = \theta_t^\rho e^{\sigma \epsilon_{t+1}}$ where we set $\rho = 1$ and $\sigma = 0.0063$. For the human capital production function we used $h_{t+1} = a_0 h_t^\rho (1 + a_1 e^\gamma)$ with $a_0 = 0.6$ and $a_1 = 1.5$. We used these parameters to emphasize

[5] The 40-period is coming soon.

the life-cycle profile of the economy.

7.1 Imperfect information over exogenous labor productivity

The left panel of Figure 3 plots the mean life-cycle profile of labor and education: education decreases over time, while hours worked increase. The right panel plots the mean productivity, that is increasing in the beginning of the life-cycle, and decreasing towards the end. As mentioned earlier with private information, this pattern generates a decreasing Frisch-elasticity, a key element for taxation. Then, as expected, mean labor taxes are increasing over the life-cycle, as seen in the left panel of Figure 4. On the other hand, mean capital taxes are decreasing. The intuition is similar to Farhi and Werning (2011).

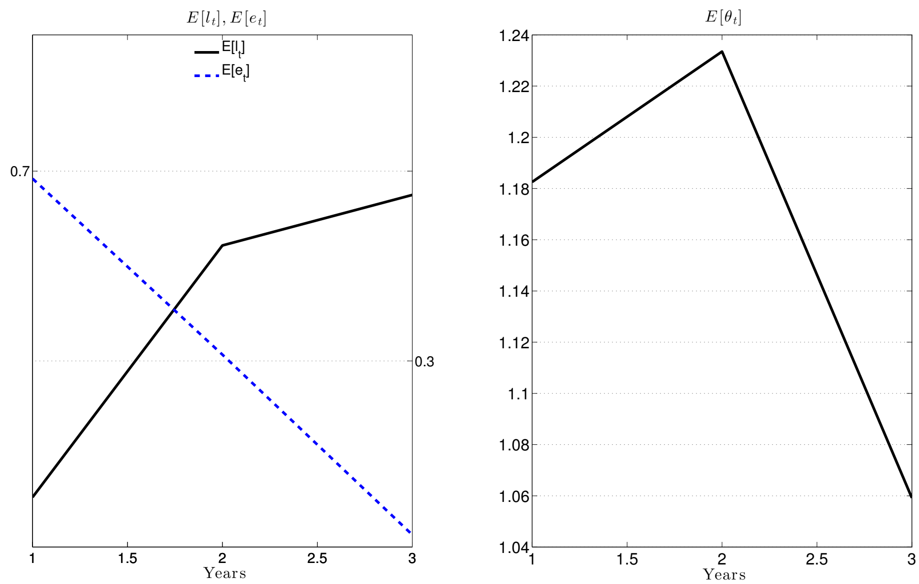

Figure 3: Life Cycle Profile of Hours Worked, Education, Productivity

Finally, we check that the first-order approach is valid in Figure 5: given the optimal contract, a worker should maximizes its utility reporting his true type: $\tilde{\theta}^r = \tilde{\theta}$. The first-order approach fails for a limited number of points whenever the policies hit the boundaries of the grid. We are working to fix it.

8 Conclusion

...

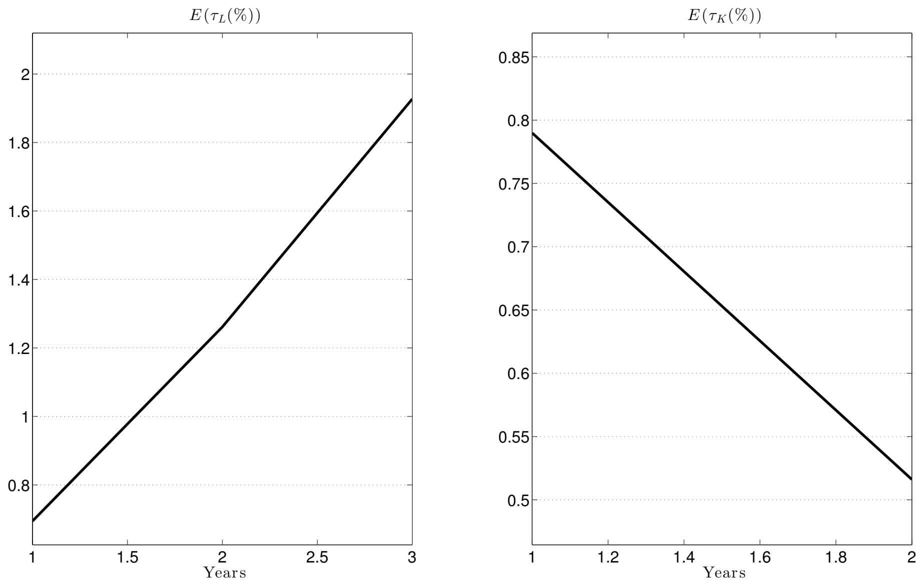

Figure 4: Life Cycle Profile of Labor and Capital Taxes

References

Bohacek, R., and M. Kapicka (2008): "Optimal Human Capital Policies," *Journal of Monetary Economics*, pp. 1–16.

Farhi, E., and I. Werning (2011): "Insurance and taxation over the life-cycle," *Working Paper*.

Golosov, M., N. R. Kocherlakota, and A. Tsyvinski (2003): "Optimal indirect and capital taxation," *The Review of Economic Studies*, 70, 569–587.

Golosov, M., M. Troshkin, and A. Tsyvinski (2011): "Optimal Dynamic Taxes," *Working Paper*.

Grochulski, B., and T. Piskorski (2010): "Risky Human Capital and Deferred Capital Income Taxation," *Journal of Economic Theoryr*, 145(3), 908–943.

Kapicka, M. (2006): "Optimal Income Taxation with Human Capital Accumulation and Limited Recordkeeping," *Review of Economic Dynamics*, pp. 612–639.

——— (2011): "The Dynamics of Optimal Taxation when Human Capital is Endogenous," *Working Paper*.

Mirrlees, J. (1971): "An exploration in the theory of optimum income taxation," *Review of Economic Studies*, 38, 175–208.

Saez, E. (2001): "Using Elasticities to Derive Optimal Income Tax Rates," *Review of Economic Studies*, 68, 205.229.

Storesletten, K., C. I. Telmer, and A. Yaron (2004): "Consumption and Risk Sharing over the Life Cycle," *Journal of Monetary Economics*, 51.

Yaron, M. H. . G. V. . A. (2011): "Sources of Lifetime Inequality," *American Economic Review*, 101(7).

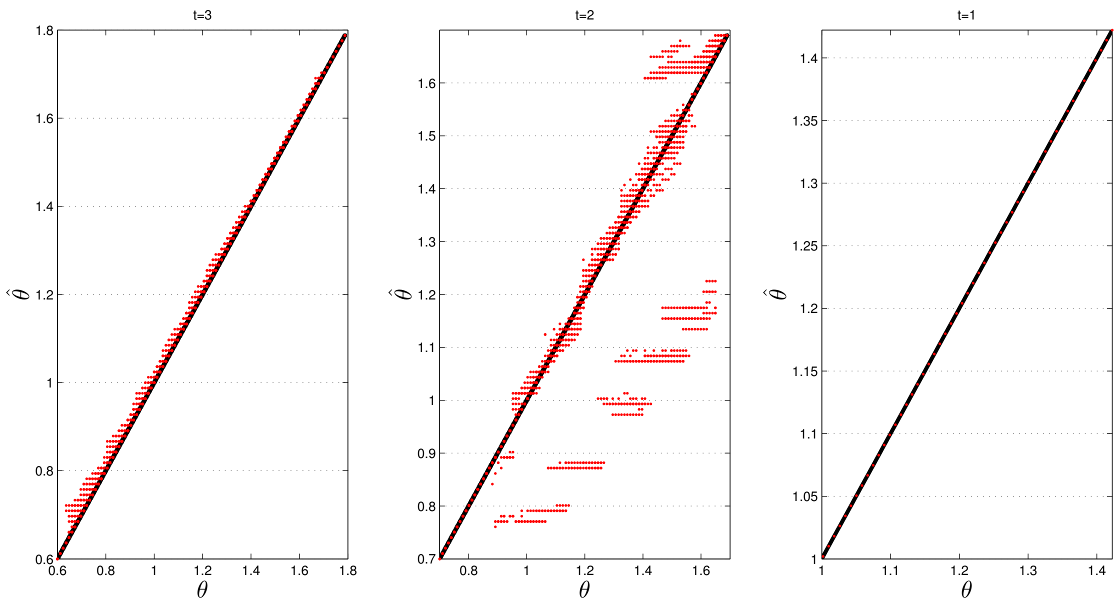

Figure 5: Numerical Check of the First-Order Approach